ZIP Centroid Matching

On the Options dialog, there are seven choices for Minimum Quality when selecting ZIP Centroid Matching. The following indicates which methodologies (the fourth character in the location code) are used for each Quality selection:

9 Digit Best Location	A, a, B, b, C, c, D, d
9 Digit Good Location	A, a, B, b, C, c, D, d, E, F, G, H
9 or 7 Digit Location	A, a, B, b, C, c, D, d, E, F, G, H, I, J, K, L, M, N
Block Group Accuracy	A, a, B, b, E, I, L
Census Tract Accuracy	A, a, B, b, E, I, L, C, c, F, J, M, V
5 Digit Only	V, W, X, Y, Z
All Centroids	Can return all Methodology codes.

Customizing the Data Locator Module

Regular Expressions

Regular expressions comprise a powerful pattern matching language. They are commonly used in the UNIX world on the command line, within text editors, and in the context of pattern matching programs such as *grep*.

A regular expression is a pattern that describes a set of ASCII strings. The fundamental building blocks are the regular expressions that match a single character. Most characters, including all letters and digits, are regular expressions that match themselves. Certain characters have special meanings. In the Centrus Data Locator module, these special characters are:
**+ * ? . # % [ ] ^ $ ( ) . **

\	"Escapes" other special characters to match the literal character.
.	Matches any character.
#	Matches a digit, same as `[0-9]`.
%	Matches a single letter, same as `[a-zA-Z]`.
~	Matches any single digit or letter, same as `[A-Za-z0-9]`
[ ]	Matches any single character of set.
^	Matches any single character not in set.
( )	Groups regular expressions together into subexpressions that can be treated as a single unit.
*	x* matches any number of x's.
+	x+ matches 1 or more x's.
?	x? matches 0 or 1 x's, where x is any single character.
|	Boolean OR character; matches one of a choice of expressions.
`^string`	Anchors match to the beginning of a string.
`string$`	Anchors match to the end of a string.

These characters can be combined in various ways to broadly or narrowly specify the strings to be matched. For example:

E-Mail	`+@(.com|COM|edu|EDU|gov|GOV|org|ORG|net|NET)`

Phone Number	(1[-]?| ? ([\(]?###[\)-]?| ?)###[-]?###(#|#?[extEXT\#] ?#+)
Social Security Number	###[-]##[-]####

More About Regular Expressions

The Centrus Data Locator module uses and extends the standard UNIX regular expression set, adding special uses for the following three characters:

- The hash or pound character (**#**) matches any single digit.
- The percent sign (**%**) matches any single letter.
- The tilde (**~**) matches any single digit or letter. Non-alphanumeric characters, such as punctuation, symbols, or ASCII "control" characters, are *not* matched.

A backslash (****) followed by any special character "escapes" the special character, matching the literal character itself. For example, to match the pound character (#), you'd use the expression **\#**.

The period (**.**) matches any single ASCII character. Thus, **.umpty** matches either "Humpty" or "Dumpty."

A range of ASCII characters may be specified by giving the first and last characters, separated by a hyphen (**-**).

A set of characters enclosed within square brackets (**[** and **]**) matches any single character in that list. For example, the regular expression **[0123456789]** matches any single digit. If the first character of the list is the caret (**^**), it matches any character *except* those in the set. It does *not* match an empty string. The caret loses this special meaning if it is not the first character of a set enclosed by brackets.

The Boolean OR character (**|**) allows a choice between two regular expressions. For example, **jell(y|ies)** matches either "jelly" or "jellies".

The following expressions are equivalent:

```
0|1|2|3|4|5|6|7|8|9
[0123456789]
[11987654321111011]
0-9
```

Parentheses (**(** **)**) group parts of regular expressions together into subexpressions that can be treated as a single unit. For example, **(ha)+** matches one or more instances of "ha".

A regular expression matching a single character may be followed by one of several repetition operators: **?**, *****, **+**. These work as follows:

string?
The *string* element is optional and matched no more than once. Thus, **xy?z** matches either "xyz" or "xz".

*string**
The *string* element will be matched zero or more times.

string+
The *string* element will be matched one or more times.

The following two expressions are equivalent:

```
[0-9]+
[0-9][0-9]*
```

All or part of a regular expression can be "anchored" to either the beginning or end of the search string. A regular expression that starts with a caret (**^**) will only match if it occurs at the beginning of the input string. A regular expression that ends with dollar (**$**) only matches at the end of the input string. Combining these two anchors constrains a regular expression to match the entire input string.

Custom Dictionaries In Data Locator

The Data Locator comes with an extensive dictionary file. This file contains:

- 48,300 first names
- 11,400 first name "aliases"
- 312,000 last names
- 36,700 company names
- 43 common name prefixes
- gender scores for first names
- ethnicity assignments for first and last names

You can define additional dictionary items in a *custom dictionary file*. Once you specify the name and location of this file, **Data locator** adds your custom dictionary entries to its own list of terms and parses input accordingly.

Creating A Custom Dictionary File

To create a custom dictionary file, you must first define the dictionary entries in a comma-delimited ASCII text file where each line is of the general format:

Type, Word[,] [Parameter,] [Parameter,] [Parameter]

Type must be one of the following: **FirstName**, **LastName**, **NickName**, **Title**, **CompanyWord**, **CompanyName**, **CityAlias**, **StateName**, **Suppress**, and **Junk**. The type is immediately followed by the word to be identified. Words are not case sensitive, and they may not have spaces

except as noted. Depending on the type, the word may be followed by one or more parameters. These are explained below.

Once you've defined the dictionary entries and saved them as an ASCII text file, use the **bldcstm.exe** program to compile the text file into an InfoStan dictionary. This command-line utility is located in your Centrus program directory, and has the following usage:

Bldcstm Infile Outfile Logfile

where:

Bldcstm starts the program
`Infile` is the name of your ASCII text source file
`Outfile` is the name to assign to the custom dictionary file, and
`Logfile` is the name to assign to the program's processing log.

Figure 2: Custom dictionary types

Type and syntax	Description
FirstName, Word, [GenderScore], [Ethnicity1], [Ethnicity2]	**FirstName** types may list a gender score of 0-100, where 0 represents a name that is feminine in 100% of cases, and 100 represents a name that is masculine in 100% of cases. The gender score may be followed by up to two ethnic identifications. (Valid ethnicities are listed in Figure 3.)
LastName, Word, [Ethnicity1], [Ethnicity2]	**LastName** types may be followed by up to two ethnic identifications. (Valid ethnicities are listed in Figure 3.)
NickName, Word, RootName	Adds a nickname to **Data locator**'s global list of nicknames.
Title, Word	Adds a title to **Data locator**'s global list of titles.
CompanyWord, Word	Indicates words that are likely to indicate the name of a business, for example, "Ltd." or "Restaurant".
CompanyName, Word	Used to add the names of specific companies to **Data locator**'s list of known company names. Spaces are permitted.
CityAlias, Word RealName	Spaces are permitted in the real name but not in the word (alias). For example, "LA" is an acceptable alias word for "Los Angeles", but "L A" is not permitted.
StateName, Word, Abbrev	A state name or variant state abbreviation is followed by the standard abbreviation. Spaces are permitted in the word. Abbreviations must be in the standard two-letter USPS form.
Suppress, Word	"Suppressed" words are dropped from all

Type and syntax	Description
	output. Spaces are permitted in the word.
Junk, `Word`	"Junk" words are words that are not assigned to any of the standard parsing categories (First Name, Last Name, Title, etc.) but can still be returned. Spaces are permitted.

Figure 3: Valid ethnicities

African	Arabic	Armenian
Asian	Basque	Belgian
Brazilian	Bulgarian	Croatian
Czech	Danish	Dutch
EastEurope	EastIndian	English
Estonian	Finnish	French
German	Greek	Hawaiian
Hungarian	Hungary	Icelandic
Iranian	Irish	Italian
Japanese	Lithuanian	NativeAmerican
Norwegeian	Pakistani	Persian
Polish	Portuguese	Rumanian
Russian	Scandinavian	Scottish
Serbian	Slavic	Hispanic
Swedish	Swiss	Turkish
Ukrainian	Vietnamese	Welsh
Many		

Examples Of Custom Dictionary Entries

```
FirstName, Grit, 0, English
LastName, Vanbiesbrouck, German, English
NickName, Grit, Margaret
Title, Representative
Title, Secretary
CompanyWord, Software
CompanyWord, Restaurant
CompanyName, Qualitative Marketing Software
CityAlias, laf, Lafayette
StateName, fla, FL
Suppress, darn
Junk, NA
```

Premium Demographics

Claritas Demographics

Current Year Estimated Population
Current Year Estimated Total Female Population
Current Year Estimated Housing Units
Current Year Estimated Median Age Total Population
Current Year Estimated Median Age Female Population
Current Year Estimated Median Age Male Population
Current Year Estimated Households
Current Year Estimated Families
Current Year Estimated Average Household Size
Current Year Estimated Percentage Of Householders Between The Age Of 15 & 24
Current Year Estimated Percentage Of Householders Between The Age Of 25 & 34
Current Year Estimated Percentage Of Householders Between The Age Of 35 & 44
Current Year Estimated Percentage Of Householders Between The Age Of 45 & 54
Current Year Estimated Percentage Of Householders Between The Age Of 55 & 64
Current Year Estimated Percentage Of Householders Between The Age Of 65 & 74
Current Year Estimated Percentage Of Householders Age 75 And Over
Current Year Estimated Average Household Income
Current Year Estimated Median Household Income
Current Year Estimated Median Family Income
Current Year Estimated Percentage Of Householders With Income Between $0 & $10,000
Current Year Estimated Percentage Of Householders With Income Between $10,000 & $15,000
Current Year Estimated Percentage Of Householders With Income Between $15,000 & $25,000
Current Year Estimated Percentage Of Householders With Income Between $25,000 & $35,000
Current Year Estimated Percentage Of Householders With Income Between $35,000 & $50,000
Current Year Estimated Percentage Of Householders With Income Between $50,000 & $75,000
Current Year Estimated Percentage Of Householders With Income Between $75,000 & $100,000
Current Year Estimated Percentage Of Householders With Income Over $100,000
Current Year Estimated Percentage Of Householders With Income Under $35,000

Current Year Estimated Percentage Of Householders With Income Over $50,000
Current Year Estimated Median Value Owner-Occupied Housing Units
Current Year Estimated Population In Group Quarters
Current Year Estimated Percentage Of White Population
Current Year Estimated Percentage Of Black Population
Current Year Estimated Percentage Of Asian Population
Current Year Estimated Percentage Of Other Population
Current Year Estimated Percentage Of Hispanic Population
Current Year Estimated Avg Hhld Expend: Grocery Stores
Current Year Estimated Avg Hhld Expend: Eating Places
Current Year Estimated Avg Hhld Expend: Drinking Places
Current Year Estimated Avg Hhld Expend: Drug Stores (Including Proprietary Stores)
Current Year Estimated Avg Hhld Expend: Building Materials And Supply Stores
Current Year Estimated Avg Hhld Expend: Hardware Stores
Current Year Estimated Avg Hhld Expend: Retail Nurseries, Lawn & Garden Supply Stores
Current Year Estimated Avg Hhld Expend: Furniture And Home Furnishing Stores
Current Year Estimated Avg Hhld Expend: Household Appliance Stores
Current Year Estimated Avg Hhld Expend: Radio, Television, And Music Stores
Current Year Estimated Avg Hhld Expend: Department Stores
Current Year Estimated Avg Hhld Expend: Apparel Stores
Current Year Estimated Avg Hhld Expend: Shoe Stores
Current Year Estimated Avg Hhld Expend: Woman's Accessory Stores
Current Year Estimated Avg Hhld Expend: Auto Supply Stores
Current Year Estimated Avg Hhld Expend: Gasoline Service Stations
Current Year Estimated Avg Hhld Expend: Jewelry Stores
Current Year Estimated Avg Hhld Expend: Book Stores
Current Year Estimated Avg Hhld Expend: Stationary Stores
Current Year Estimated Avg Hhld Expend: Gift, Novelty, And Souvenir Shops
Current Year Estimated Avg Hhld Expend: Florist
Current Year Estimated Avg Hhld Expend: Hobby, Toy, And Game Shops
Current Year Estimated Avg Hhld Expend: Sporting Goods Stores And Bicycle Shops
Current Year Estimated Avg Hhld Expend: Camera And Photographic Supply Stores
Current Year Estimated Avg Hhld Expend: Luggage Stores
Current Year Estimated Percentage Of Married Couples/Families
Current Year Estimated Percentage Of Married Couples/Families With Related Children
Current Year Estimated Percentage Of Married Couples/Families With Related Children Under Age 5
Current Year Estimated Percentage Of Married Couple/Families With Related Children Age 5-17

Current Year Estimated Percentage Of Population 25+ With Education Below 9th Grade
Current Year Estimated Percentage Of Population 25+ With Education Between 9th & 12th Grades With No Diploma
Current Year Estimated Percentage Of Population 25+ Diploma As Highest Education (Includes Equivalency)
Current Year Estimated Percentage Of Population 25+ Completing 2 Years Or Less Of College
Current Year Estimated Percentage Of Population 25+ With Bachelor's Degree As Highest Education
Current Year Estimated Percentage Of Population 25+ With Graduate Or Professional Degree
Current Year Estimated Persons In Occupied Units In Buildings With 2 To 4 U.I.S.
Current Year Estimated Persons In Occupied Housing Units With 5 To 9 Units In Structure
Current Year Estimated Persons In Occupied Units In Buildings With 10-19 U.I.S.
Current Year Estimated Persons In Occupied Units In Buildings With 20-49 U.I.S.
Current Year Estimated Persons In Occupied Housing Units with 50+ Units In Structure
Current Year Estimated Persons In Occupied Units Considered Mobile Homes Or Trailers
5-Year Projected Population
5-Year Projected Total Female Pop
5-Year Projected Housing Units
5-Year Projected Median Age Total Pop
5-Year Projected Median Age Female Pop
5-Year Projected Median Age Male Pop
5-Year Projected Households
5-Year Projected Families
5-Year Projected Average Household Size
5-Year Projected Percentage Of Householders Between The Age Of 15 & 24
5-Year Projected Percentage Of Householders Between The Age Of 25 & 34
5-Year Projected Percentage Of Householders Between The Age Of 35 & 44
5-Year Projected Percentage Of Householders Between The Age Of 45 & 54
5-Year Projected Percentage Of Householders Between The Age Of 55 & 64
5-Year Projected Percentage Of Householders Between The Age Of 65 & 74
5-Year Projected Percentage Of Householders Age 75 And Over
5-Year Projected Average Household Income
5-Year Projected Median Household Income
5-Year Projected Median Family Income
5-Year Projected Percentage Of Householders With Income Between $0 & $10,000
5-Year Projected Percentage Of Householders With Income Between $10,000 & $15,000
5-Year Projected Percentage Of Householders With Income Between $15,000 & $25,000

5-Year Projected Percentage Of Householders With Income Between $25,000 & $35,000
5-Year Projected Percentage Of Householders With Income Between $35,000 & $50,000
5-Year Projected Percentage Of Householders With Income Between $50,000 & $75,000
5-Year Projected Percentage Of Householders With Income Between $75,000 & $100,00
5-Year Projected Percentage Of Householders With Income Over $100,000
5-Year Projected Percentage Of Householders With Income Under $35,000
5-Year Projected Percentage Of Householders With Income Over $50,000
5-Year Projected Median Value Owner-Occupied HUS
5-Year Projected Population In Group Quarters
5-Year Projected Percentage Of White Population
5-Year Projected Percentage Of Black Population
5-Year Projected Percentage Of Asian Population
5-Year Projected Percentage Of Other Population
5-Year Projected Percentage Of Hispanic Population
5-Year Projected Avg Hhld Expend: Grocery Stores
5-Year Projected Avg Hhld Expend: Eating Places
5-Year Projected Avg Hhld Expend: Drinking Places
5-Year Projected Avg Hhld Expend: Drug Stores (Including Proprietary Stores)
5-Year Projected Avg Hhld Expend: Building Materials And Supply Stores
5-Year Projected Avg Hhld Expend: Hardware Stores
5-Year Projected Avg Hhld Expend: Retail Nurseries, Lawn & Garden Supply Stores
5-Year Projected Avg Hhld Expend: Furniture And Home furnishing Stores
5-Year Projected Avg Hhld Expend: Household Appliance Stores
5-Year Projected Avg Hhld Expend: Radio, Television, And Music Stores
5-Year Projected Avg Hhld Expend: Department Stores
5-Year Projected Avg Hhld Expend: Apparel Stores
5-Year Projected Avg Hhld Expend: Shoe Stores
5-Year Projected Avg Hhld Expend: Woman's Accessory Stores
5-Year Projected Avg Hhld Expend: Auto Supply Stores
5-Year Projected Avg Hhld Expend: Gasoline Service Stations
5-Year Projected Avg Hhld Expend: Jewelry Stores
5-Year Projected Avg Hhld Expend: Book Stores
5-Year Projected Avg Hhld Expend: Stationary Stores
5-Year Projected Avg Hhld Expend: Gift, Novelty, And Souvenir Shops
5-Year Projected Avg Hhld Expend: Florists
5-Year Projected Avg Hhld Expend: Hobby, Toy, And Game Shops
5-Year Projected Avg Hhld Expend: Sporting Goods Stores And Bicycle Shops
5-Year Projected Avg Hhld Expend: Camera And Photographic Supply Stores
5-Year Projected Avg Hhld Expend: Luggage Stores
1990 Occupied Single, Detached Housing Units
1990 Occupied Single, Attached Housing Units
Housing Units Built 1985-1990

Housing Units Built 1980-1984
Housing Units Built 1970-1979
Housing Units Built Before 1970
Occupied Housing Units, Hhldr Moved In 1989 To March 1990
Occupied Housing Units, Hhldr Moved In 1985-1988
Occupied Housing Units, Hhldr Moved In 1980-1984
Occupied Housing Units, Hhldr Moved In Before 1980
Percentage Of Population Commuting To Work In 10 Minutes Or Less
Percentage Of Population Commuting To Work Between 10 And 19 Minutes
Percentage Of Population Commuting To Work Between 20 And 29 Minutes
Percentage Of Population Commuting To Work In 30 Or More Minutes
Percentage Of Households With No Vehicle
Percentage Of Households With One Vehicle
Percentage Of Households With 2 Or More Vehicles

National Decision Systems Demographics

Current Year Estimated Population (Estimated)
Current Year Estimated Pop By Household Type: Group Quarters
Current Year Estimated Households
Current Year Estimated Average Household Size
Current Year Estimated Total Male Population
Current Year Estimated Total Female Population
Current Year Estimated Median Age Total Population
Current Year Estimated Median Age Male Population
Current Year Estimated Median Age Female Population
Current Year Estimated % Of Households With Income $15k - $25k
Current Year Estimated % Of Households With Income $25k - $35k
Current Year Estimated % Of Households With Income $35k - $50k
Current Year Estimated % Of Households With Income $50k - $75k
Current Year Estimated % Of Households With Income $75k - $100k
Current Year Estimated % Of Households With Income $100k - $150k
Current Year Estimated % Of Households With Income $150k+
Current Year Estimated Average Household Income
Current Year Estimated Median Household Income
Current Year Estimated % Of White Population
Current Year Estimated % Of Black Population
Current Year Estimated % Of Asian, Pacific Island Pop
Current Year Estimated % Of American Indian And Other Pop
Current Year Estimated % Of Hispanic Population
Current Year Estimated % Of Households Age 0 - 24
Current Year Estimated % Of Households Age 25 - 34
Current Year Estimated % Of Households Age 35 - 44
Current Year Estimated % Of Households Age 45 - 54
Current Year Estimated % Of Households Age 55 - 64
Current Year Estimated % Of Households Age 65 - 74
Current Year Estimated % Of Households Age 75+
Current Year Estimated Sales: Building Materials, Hardware, Garden
Current Year Estimated Sales: Motor Vehicle, Misc. Auto Dealer
Current Year Estimated Sales: Automotive, Home Supply Stores

Current Year Estimated Sales: Furniture, Home Furnishings
Current Year Estimated Sales: Household Appliance, Radio, TV, Stereos
Current Year Estimated Sales: General Merchandise Group
Current Year Estimated Sales: Department Stores, Leased Department
Current Year Estimated Sales: Food Stores
Current Year Estimated Sales: Gasoline Service Stations
Current Year Estimated Sales: Apparel Accessory Stores
Current Year Estimated Sales: Eating, Drinking Places
Current Year Estimated Sales: Drug Proprietary Stores
Current Year Estimated Sales: Total Retail Trade
5-Year Projected Population (Projected)
5-Year Projected Pop By Household Type: Group Quarters
5-Year Projected Households
5-Year Projected Average Household Size
5-Year Projected Total Male Population
5-Year Projected Total Female Population
5-Year Projected Median Age Total Population
5-Year Projected Median Age Male Population
5-Year Projected Median Age Female Population
5-Year Projected % Of Households With Income $15k - $25k
5-Year Projected % Of Households With Income $25k - $35k
5-Year Projected % Of Households With Income $35k - $50k
5-Year Projected % Of Households With Income $50k - $75k
5-Year Projected % Of Households With Income $75k - $100k
5-Year Projected % Of Households With Income $100k- $150k
5-Year Projected % Of Households With Income $150k+
5-Year Projected Average Household Income
5-Year Projected Median Household Income
5-Year Projected % Of White Population
5-Year Projected % Of Black Population
5-Year Projected % Of Asian, Pacific Island Pop
5-Year Projected % Of American Indian And Other Pop
5-Year Projected % Of Hispanic Population
5-Year Projected % Of Households Age 0 - 24
5-Year Projected % Of Households Age 25 - 34
5-Year Projected % Of Households Age 35 - 44
5-Year Projected % Of Households Age 45 - 54
5-Year Projected % Of Households Age 55 - 64
5-Year Projected % Of Households Age 65 - 74
5-Year Projected % Of Households Age 75+
1990 % Of Married Population
1990 % Of Households With Married Couples
1990 % Of Households With Married Couples & Child
1990 Owner Occupied Housing
1990 % Of Owner Occupied Housing Unit Val $0k - $25k
1990 % Of Owner Occupied Housing Unit Val $25k - $50k
1990 % Of Owner Occupied Housing Unit Val $50k - $75k
1990 % Of Owner Occupied Housing Unit Val $75k - $100k
1990 % Of Owner Occupied Housing Unit Val $100k - $150k
1990 % Of Owner Occupied Housing Unit Val $150k - $200k

1990 % Of Owner Occupied Housing Unit Val $200k - $300k
1990 % Of Owner Occupied Housing Unit Val $300k - $400k
1990 % Of Owner Occupied Housing Unit Val $400k - $500k
1990 % Of Owner Occupied Housing Unit Val $500k+
1990 Median Val Owner Occupied Housing Units
1990 Occupied Single, Detached Housing Units
1990 Occupied Single, Attached Housing Units
1990 Occupied Housing 2 Units In Structure
1990 Occupied Housing 3-9 Units In Structure
1990 Occupied Housing 10-19 Units In Structure
1990 Occupied Housing 20-49 Units In Structure
1990 Occupied Housing 50+ Units In Structure
1990 Occupied Housing Considered Mobile
1990 Occupied Housing Considered Other
1990 % Of Pop Commute To Work 0-9 Min
1990 % Of Pop Commute To Work 10-29 Min
1990 % Of Pop Commute To Work 30-59 Min
1990 % Of Pop Commute To Work 60-89 Min
1990 % Of Pop Commute To Work 90+
1990 Average Commute To Work
1990 % Of Pop Education Below 9th Grade
1990 % Of Pop Education 9th & 12th Grades
1990 % Of Pop High School Diploma
1990 % Of Pop Less Than 2 Years Of College
1990 % Of Pop With Associates Degree
1990 % Of Pop With BS Degree
1990 % Of Pop With Grad Or Professional Degree
1990 % Of Housing Units Built 1989-1990
1990 % Of Housing Units Built 1985-1988
1990 % Of Housing Units Built 1980-1984
1990 % Of Housing Units Built 1970-1979
1990 % Of Housing Units Built 1960-1969
1990 % Of Housing Units Built 1950-1959
1990 % Of Housing Units Built 1940-1949
1990 % Of Housing Units Built 1939 Or Earlier

GSD Split for Windows

About GSD Split For Windows

GSD Split for Windows can be used to split US.GSD, UST.GSD and US.Z9 files into smaller files. This is commonly done to speed processing when only a specific geography is needed. GSD Split divides GSD and Z9 files by ZIP Code, SCC (3-digit ZIP), MSA or State, and remembers the last extract performed, which makes extracting the updates an easier process.

GSD (both primary and supplemental) and Z9 files are used by a variety of Qualitative Marketing software. These files contain the address standardization and geocoding data used by GeoStan, Centrus and StarData products. Using GSD Split, you can split all of these files in one pass. If you are using the supplemental files and have a single CD drive, GSD Split will prompt you to place that CD in the drive at the proper time.

You can use GSD Split to split files as needed, but be aware of the following requirements:

- If you are using supplemental file data, be sure that the supplemental file is extracted using the same geography as the primary file. The easiest way to insure this is to extract the primary and supplemental files at the same time.
- If you are using any external indexes, which are in a GSI file (such as State-wide Intersection, Enhanced Alias, etc.,) you cannot use GSD Split. Please contact Qualitative Marketing Software if you need to split files for use with these indices.

Note: GSD Split experiences file manipulation inconsistencies when reading from and writing to Novell network drives.When splitting on Novell networks, process to and from your local machine only.After GSD Split operations are complete, you may move the split files to a network drive. No problems have been observed on any other networks.

Installing GSD Split For Windows

GSD Split for Windows is installed automatically when you install any Qualitative Marketing software applications that use the GSD or Z9 files. For information on these programs, please refer to the installation notes that came with the CD

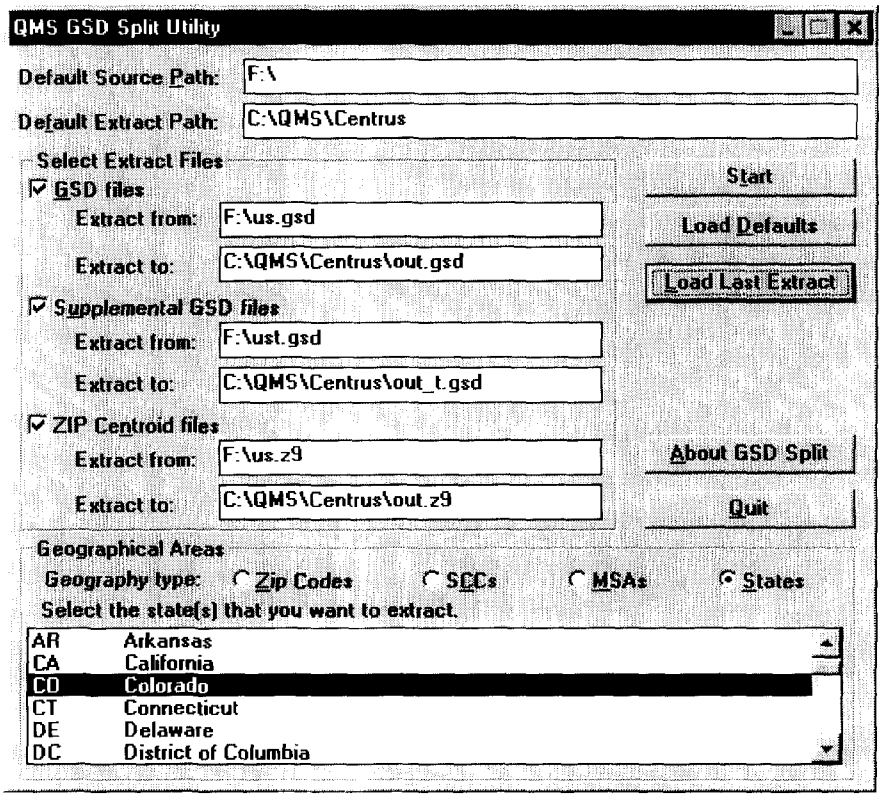
Starting GSD Split For Windows

To start GSD Split For Windows, choose **GSD Split** from the Centrus program group on the Windows Start menu.

Using GSD Split For Windows

The GSD Split Utility main dialog box shown below is configured to extract the state of Colorado from the primary and supplemental GSD files, as well as from the Z9 file. The files will be extracted from the F: drive (the CD drive) to a directory called C:\QMS\Centrus. Each dialog option is discussed in detail below.

Figure 4: GSD Split Utility dialog box



Default Path Settings

Default Source Path—This is the default path where GSD Split will look for the required support files (CTYST.DIR and MSABYZIP.TXT) as well as for the GSD and Z9 files, unless otherwise specified in the Select Extract Files section of the dialog. In almost all cases, it will be the CD drive where the Qualitative Marketing Software CD is located. GSD Split will prompt you to swap CDs if you are extracting both primary and supplemental data.

Default Extract Path—This path will be used as the default path to extract to, unless otherwise specified in the Select Extract Files section of the dialog.

Select Extract Files Section

In this section, specific file names and paths may be given for all source and extract files. The default uses "OUT" for the extract names, and the default paths are entered in the top section of the dialog.

GSD Files—If this box is checked, the primary GSD file is extracted.

Supplemental GSD Files—If this box is checked, the supplemental GSD file is extracted.

ZIP Centroid Files—If this box is checked, the Z9 file is extracted.

Extract From—For each file type (Primary, Supplemental and Z9), the Extract From argument should be filled with the path and name to the original file (most often, US.GSD, UST.GSD and US.Z9).

Extract To—For each file type (Primary, Supplemental and Z9), the Extract To argument should be filled with the path and name to the extract file to create. Filename extensions (GSD and Z9) must be preserved for these files to work correctly, but the files may be named anything you like.

Geographical Areas Section

This section allows you to specify the geographic areas that should be extracted. You may specify the area by ZIP Code, SCC (3-digit ZIP Codes), MSAs or States.

Selecting by ZIP—To extract by ZIP, type in the ZIP Code range, using hyphens, or enter discrete ZIP Codes, delimited with commas. For example, *80301-80503, 81456* extracts all areas that contain ZIP Codes between 80301 and 80503, and the single 81456 ZIP Code.

Selecting by SCC—To extract by SCC, which are also referred to as 3-digit ZIPs, type in the SCC range, using hyphens, or enter discrete SCC Codes, delimited with commas. For example, *803-805, 814* extracts all areas that contain SCC Codes between 803 and 805, and the single 814 SCC Code.

Selecting by MSA—To extract by MSA, highlight the MSAs (using the space bar or left mouse key) you wish to extract.

Selecting by State—To extract by State, highlight the States (using the space bar or left mouse key) you wish to extract.

GSD Extract Buttons

Start—Pressing this button begins the extraction. Depending upon the speed of your machine and CD ROM drive, extraction can take 15 minutes or more for large geographic areas

Load Defaults—Pressing this button loads the default information, which normally consists of the Default Source and Extract Paths. If the CD containing the GSD or Z9 files is not in the appropriate drive, GSD Split may not display any defaults.

Load Last Extract—Pressing this button loads the last extract performed. This is very handy if you want to perform the same extract for each release update.

About GSD Split—Brings up the "About" box for GSD Split.

Quit—Exits the GSD Split program.

Preparing ASCII Text Files

About Format Files

While most files can be immediately used within Centrus, ASCII files require a special setup file, called a *format file*. Centrus comes with a utility program called FMTmaker which automates the process of creating and updating format files. FMTmaker reads field and record information from the existing format file, if there is one, or from the actual ASCII text file. This information is then displayed, along with a section of the first part of the text file so that you can see how the data will be interpreted by Centrus. You can accept FMTmaker's settings or make changes as necessary. When you're satisfied that your data is correctly described, FMTmaker will update any existing format file or create a new one. Your ASCII text file is then ready for processing by Centrus or StarData.define data fields using the FMTMaker utility.

The format file describes the organization of the data in the ASCII file. Centrus will accept three types of ASCII text file:

- Files with fixed field length and an end of line character at the end of each record.
- Files with fixed field length and without an end of line character.
- Files with comma or tab delimiters and an end of line character at the end of each record.

Centrus will use the settings defined in the format file to read the ASCII text file. During processing, Centrus will create a new ASCII file using the original filename. The old file will be renamed with the extension **.BAK**.

The format file must have the same file name as the ASCII text file but use the extension **.FMT** rather than **.TXT**. Blank lines are ignored in the format file, as are comments prefaced by slashes (/) or semicolons (;). The file may be in upper, lower or mixed case.

The beginning of the file can contain three tokens: `TYPE`, for type of delimiter; `EOL`, for end of line indicator; and `FLN`, to indicate whether field names appear in the first line.

- Delimiter `TYPE` can be Fixed, Comma or Tab.
- `EOL` can be `CRLF`, Return, Linefeed or None.
- If field names appear on the first line of a **delimited file**, set `FLN=y`. The first line will then be ignored during processing. If field names do not appear on the first line *or* if it is a **fixed length file**, set `FLN=n`.

Note: If these tokens are not described they will be assigned the following default values:

```
Type: Fixed
EOL: CRLF
FLN: N
```

In addition to these three tokens, the actual fields within the ACSII text file are described. All fields must be listed in order of occurrence, regardless of whether they will be accessed by Centrus . The list of fields must fully describe the entire input record

Each field name must be followed by a field type. If the file is fixed length, field size is also required. and numeric fields may have an optional number of decimal places. The name, type and size must be separated by tabs or spaces, and the number of decimal places (if present) must be separated from the size by a period.

- Field names must be ten characters or less in length to be completely visible in Centrus. Field names must be unique and cannot contain blank spaces.
- Field types are either character or numeric, which may be abbreviated to "C" and "N".

If the EOL is not "None," you may add fields at the end of the .FMT file that are not in the original ASCII list file.

Delimited File Example

A sample format file for a delimited text file is shown below:

```
; Comments begin with semicolons (;) or slashes (/).
; Size is not required as this is a comma delimited file.
TYPE=Comma
EOL=CRLF
FLN=N
First_Name Character
Last Name  Character
Address Character
City    Character
State   Character
ZIP Character
Latitude    Numeric
Longitude   Numeric
Block_Grp   Character
Match_Cd    Character
Loc_Code    Character
```

Sample data records are shown below. Note that all fields are not present for each record, but blank fields between two existing fields are still delimited. Note, too, that each field is enclosed by quotation marks indicating literal strings. This prevents any commas that may be present within fields from being used as delimiters. Latitude, Longitude, Block_Grp, Match_Cd and Loc_Code fields will all be added to the file during processing.

```
"Bob","Doe","123 Main St","Nome","AK"
"Carol","Jones","321 28th Ave, Apt. 22","","","80301"
"Ted","Tyler","345 Mission Cir","Longmont","CO","80503"
"Alice","Brown","P.O. Box 12","Boulder","CO"
```

Fixed Length File Example

A sample format file for a fixed length text file is shown below:

```
; Comments begin with semicolons (;) or slashes (/).
; Field size IS required! This is a fixed length file.
TYPE=Fixed
EOL=CRLF
FLN=n
First_Name Character  1
Last_Name  Character  1
Address Character  2
City    Character  1
State   Character  2
ZIP Character  5
Latitude    Numeric 11 6
Longitude   Numeric 11 6
Block_Grp   Character  1
Match_Cd    Character  1
Loc_Code    Character  1
```

Sample data records are shown below. Note that all fields are not present for each record, but blank fields between two existing fields are filled with blank spaces. Latitude, Longitude, Block_Grp, Match_Cd and Loc_Code fields will all be added to the file during processing.

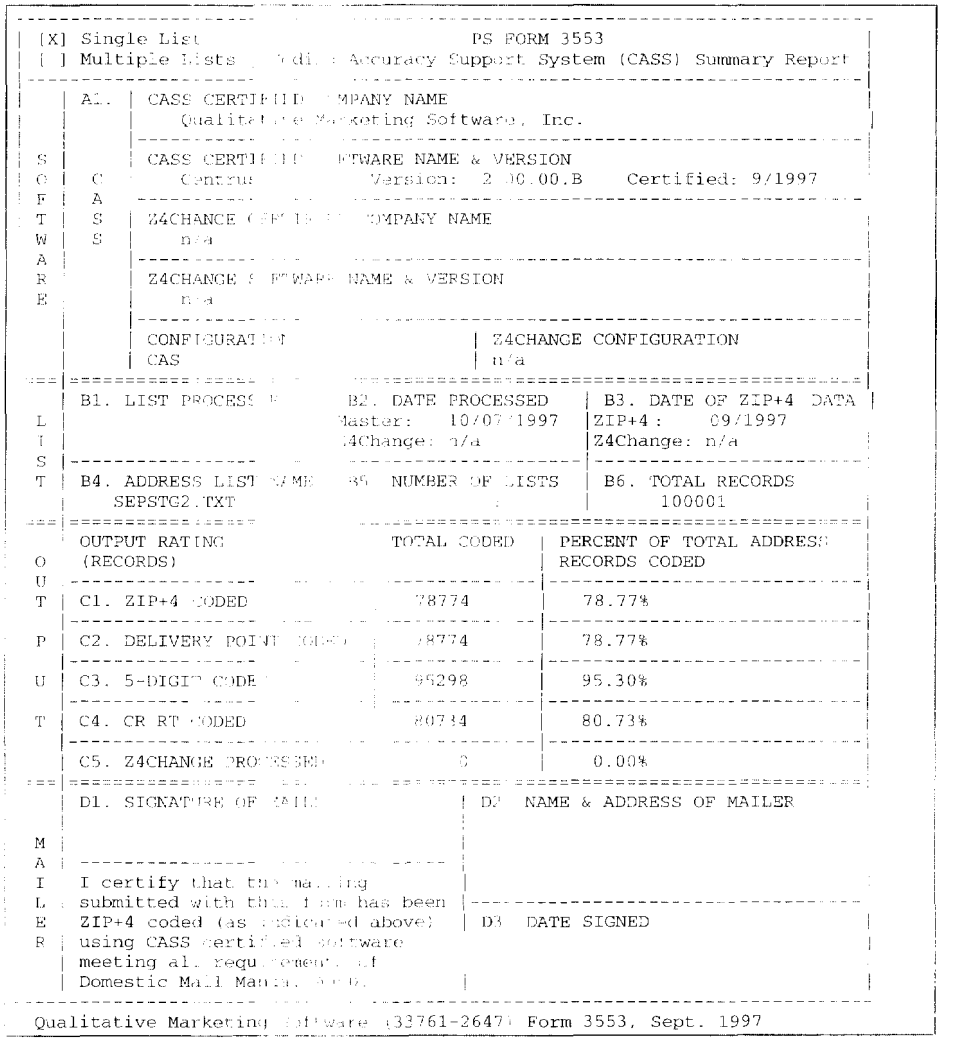
```
Bob         Doe         23 Main St     Nome        AK
Carol       Jones       21 18th Ave                    80301
Ted         Tyler       45 Mission St  Longmont    CO 80503
Alice       Brown       10 BOX 123     Boulder     CO
```

CASS Report and Bulk Rate Information

Cass (3553) Report Information

The report generated by Centrus Desktop is the USPS 3553 report, commonly referred to as the CASS report. When a CASS report is requested, Centrus compiles statistics during processing regarding the number of 5-digit ZIP Codes, ZIP+4 Codes and Carrier Routes assigned to records in the address file. It prints this information, along with other information required by the USPS, in an ASCII text file. When processing is complete, Centrus Desktop displays the report file in Windows Notepad, if available.

An example of the CASS Report is shown below:



Before submitting this report, you must supply information in several sections. Section B1 must be filled in with the company that processed the list. Section D2 contains the company that is doing the mailing. Sections B1 and D2 are often the same. The company doing the mailing must also sign in